BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

ESCHERICHIA COLI O157:H7 EPITHELIAL ADHESIN

Field of the Invention

The invention relates to genetic engineering and particularly to the demonstration that a contiguous segment of chromosomal DNA from *E. coli* O157:H7 encodes an adhesin that mediates colonization of the gastrointestinal tracts of bovines, and possibly humans, with *E. coli* O157:H7 and bacteria using structurally related adherence mechanisms.

Background of the Invention

E. coli O157:H7 is a virulent and common foodborne pathogen. Most outbreaks, and many sporadic cases (38,42; see the appended Citations), have been attributed to food of bovine origin. Most *E. coli* O157:H7 infections are sporadic, but this organism can cause massive epidemics by contamination of ground beef (19) and water (69). *E. coli* O157:H7 is transmissible from person to person, but the disappearance of the strain which caused the massive 1993 outbreak in Washington State soon after recall of the incriminated vehicle demonstrates that ingestion of contaminated beef, and not person to person spread, is the chief source of human infection.

E. coli O157:H7 organism elaborates Shiga-like toxins (SLT) I and/or II. SLT I and II inhibit protein synthesis by disrupting a glycosidic bond at a specific adenine (A4324) in 28S rRNA of the 60S ribosomal subunit. SLT-producing *E. coli* (SLTEC) are ubiquitous in food (62) and animals (47). The vast majority are probably not human pathogens.

Current data suggest that *E. coli* O157:H7 is the most common and medically significant SLTEC. Only one outbreak of bloody diarrhea caused by SLTEC other than *E. coli* O157:H7 has ever been reported (11). Additionally, even when sought appropriately, non-O157:H7 SLTEC are rarely found in stools submitted for bacterial culture in North America compared to their frequency in the environment (8,52,59). Moreover, *E. coli*

O157:H7 is the predominant precipitant of the hemolytic uremic syndrome (HUS), the most important complication of enteric infection with *E. coli* O157:H7. For example, *E. coli* O157:H7 was found in 96% of HUS patients if stool was obtained within the first six days of diarrhea (72). Even though non-O157:H7 SLTEC have caused some cases of HUS in several foreign series (10,11,12,35,40), these strains have never been reported to cause HUS in the United States. These data suggest that *E. coli* O157:H7 is the most important cattle-borne human pathogen threatening the food supply of this country today.

Cattle are the only reservoir of *E. coli* O157:H7 so far identified. Approximately 1 in 200 apparently healthy northwestern United States dairy and beef cattle carry *E. coli* O157:H7, and 8 to 16% of herds have at least one infected animal (25). Similar carriage rates have been detected nationwide (26). These are probably minimum carriage rates, because the technique used to culture *E. coli* O157:H7 is relatively insensitive.

A very low inoculum of *E. coli* O157:H7 can cause human disease. Person to person spread occurs rather easily in outbreaks and among sporadic cases (5,6,60). Microbiologic analysis of the contaminated hamburger from the 1993 Western United States outbreak demonstrated that only approximately 200 *E. coli* O157:H7 were present in each of the contaminated patties (46). It is probable that the inadequate cooking that was applied reduced this concentration by at least one log, suggesting that very few *E. coli* O157:H7, perhaps in the range of 1-10 bacteria, can cause clinically apparent infection.

Data suggest that the incidence of diseases caused by *E. coli* O157:H7 has increased in the United States, independent of ascertainment bias by diagnosing physicians (44,70). Additionally, an increasing rate of antibiotic resistance in Washington State human isolates of *E. coli* O157:H7 might portend an increased prevalence of this pathogen in animals administered antibiotics. For example, before 1988, none of 56 strains of *E. coli* O157:H7 were resistant to a wide variety of antibiotics tested, whereas after 1988, 7.4% of 176 strains were resistant to the same combination of antimicrobials (streptomycin, sulfamethoxazole, and tetracycline). It is probable that the selective pressure for the acquisition of antibiotic resistance in *E. coli* O157:H7 occurred in farm animals. This emerging resistance is of considerable concern because such strains might achieve a selective advantage over other coliform bacilli in cattle given antibiotics, thereby increasing the frequency with which food of bovine origin is contaminated with this pathogen.

Because of the ease with which *E. coli* O157:H7 can cause human disease, it is crucial to reduce this pathogen in, or eliminate it from, its ecological niche, namely the gastrointestinal tracts of healthy cattle.

The molecular mechanisms used by *E. coli* O157:H7 to adhere to epithelial cells and colonize animals are poorly characterized. However, the adhesive properties of *E. coli* 0157:H7 have been noted by several investigators. Most North American strains of *E. coli*

O157:H7 displayed D-mannose-resistant adherence patterns to HEp-2 or Henle 407 cells (57). Most strains adhere in the form of localized microcolonies, a phenotype strongly linked to diarrhea in epidemiological studies of enteropathogenic *E. coli* (EPEC) (13,16). A 60 MDa plasmid is present in all strains of *E. coli* O157:H7, and one group associated the expression of sparse D-mannose-resistant adhesion to Henle 407 cells to the presence of this plasmid (34). Plasmid-cured *E. coli* O157:H7 expressed no fimbriae and were nonadherent, and a 60 MDa plasmid from *E. coli* O157:H7 conferred weak adherence to non-adherent *E. coli* C600. However, other investigators have shown that plasmid-less *E. coli* O157:H7 were fimbriated, whereas laboratory *E. coli* strains were not (79). Furthermore, plasmid-cured *E. coli* O157:H7 adheres to epithelial cells as well or better than its parent (22,33). Only one of five adherent strains of *E. coli* O157:H7 studied by Sherman et al. (66) was fimbriated, but this fimbriated strain also agglutinated erythrocytes. The agglutination was sensitive to D-mannose, suggesting that this adherence was due to type I fimbriae. Taken together, these data suggest that an identifiable fimbrial structure is not responsible for the adherence of most *E. coli* O157:H7 to Henle 407 cells.

Outer membranes of *E. coli* O157:H7 competitively inhibit adherence to HEp-2 cells, an inhibition which is not due to H7 flagellin or O157 lipopolysaccharide (65). Adherence of *E. coli* O157:H7 to HEp-2 cells was reduced, but not abolished, by antibody to a 94 kDa outer membrane protein (64). Antibodies to enterotoxigenic *E. coli* colonization factor antigens I and II do not detect surface structures on *E. coli* O157:H7 (78). *E. coli* O157:H7 do not have sequences homologous to the EPEC adherence factor plasmid or to the diffuse adherence adhesin (71).

Some investigators have suggested that the epithelial cell adhesin of *E. coli* O157:H7 is encoded by its *eae* gene (17). *E. coli* O157:H7 *eae* is related to *inv*, which encodes *Yersinia* invasin, which also functions as an adhesin, and EPEC *eae*, which encodes intimin. An *eae* deletion mutant of *E. coli* O157:H7 neither adhered to HEp-2 cells nor caused the attaching and effacing (AE) lesion in newborn pigs (17). When deletion mutants were complemented in *trans* by an intact *eae* gene, the strain could again cause the AE lesion, but still could not adhere *in vitro*. However, data from other groups suggest that the *eae* gene product is not an adhesin for *E. coli* O157:H7. First, despite sequence homology to *inv* in its bacterial localization and transmembrane domains, the receptor binding domain of *E. coli* O157:H7 *eae* is quite dissimilar (4,82). Second, an *eae* insertional mutant in *E. coli* O157:H7 retained the ability to adhere to HEp-2 cells in a quantitative adherence assay (41). Third, an *eae* gene product does not confer adherence on nonadherent laboratory strains of *E. coli*. (Jerse, A., et al., Proc. Natl. Acad. Sci. USA 87:7839-7843, 1990) Thus, a molecule other than the *eae* gene product in *E. coli* O157:H7 appears to be the primary adhesin of *E. coli* O157:H7 for bovine epithelial cells, enabling this human pathogen to colonize the bovine

gastrointestinal tract.

Bacterial adhesins, when used as immunogens, prevent disease or colonization of mucosal surfaces by bacteria in many animals (1,18,21,29,30,36,49,50,55,61, 68,81, which are hereby incorporated by reference). The reduction of *E. coli* O157:H7 at its bovine source would enhance the microbiologic safety of food derived from cattle, and lessen the environmental biohazard risk posed by the approximately 100,000 cattle detectably infected with *E. coli* O157:H7 at any one time in the United States. The availability of antibody for passive immunization would greatly mitigate the harm engendered by outbreaks of this infection.

Summary of the Invention

Transposon-mediated mutations of *E. coli* O157:H7 have been isolated that do not adhere to HeLa cells and that have lost the ability to colonize bovine intestines. A HeLa cell *in vitro* system has been established that provides a means of assaying variants of *E. coli* 0157:H7 for their ability to colonize cattle. The gene into which the transposon inserted have been sequenced. In a seperate approach, two overlapping 40 kb segments of chromosomal DNA from *E. coli* O157:H7 have been cloned that confer D-mannose resistant adherence to nonadherent strains of *E. coli*. The overlapping region has been cloned, and *E. coli* HB101 expressing this overlapping region on a plasmid (ATCC No. 69648) have acquired the ability to adhere to epithelial cells of both human (HeLa) and bovine (MDBK) origin. These findings demonstrate that a contiguous segment of chromosomal DNA from *E. coli* O157:H7 encodes an adhesin, and that this same adhesin mediates both bacterial adherence to HeLa cells.

The adhesin-encoding region of p*ear* has been identified as the nucleotide sequence of SEQ ID NO:4. Also described are recombinant expression vectors containing the adhesin-coding region, as are bacterial cells which are transformed with the recombinant expression vector. The recombinant adhesin preferably has the amino acid sequence of SEQ ID NO:4. Also described are immunological binding partners that bind to the recombinant adhesin. Vaccine formulations of the invention contain the recombinant adhesin encoded by a nucleic acid molecule that hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NO:4.

Brief Description of the Drawings

FIGURE 1 shows shedding rates of *E. coli* O157:H7 from calves inoculated with adherent strain 86-24 and non-adherent strain F4. Bars = SE.

FIGURE 2 is a map demonstrating overlap region between cloned regions of pSC(A-G6) and pSC(T-H12), and an adherence-conferring subclone region of the overlap, p*ear*.

Detailed Description of the Preferred Embodiment

Escherichia coli O157:H7 causes severe and potentially fatal infections in humans,

especially children. This pathogen is harbored by apparently healthy cattle. The molecular mechanisms underlying the carriage of *E. coli* O157:H7 by cattle are not understood. The invention provides an isolated adhesin that enables *E. coli* O157:H7 to adhere to epithelial cells *in vitro* and to colonize cattle.

Transposon-based mutants of *E. coli* O157:H7 have been identified that have lost the ability to adhere to epithelial cells. These same mutants colonize calves poorly compared to parental *E. coli* O157:H7. This finding validates the use of the HeLa cell adherence assay. Additionally, a region of the *E. coli* O157:H7 chromosome has been cloned that expresses the adhesin. This chromosomal region has been designated the epithelial adherence region (ear).

Two complementary approaches were used to identify the gene encoding the adhesin that enables *E. coli* O157:H7 to colonize bovine intestines. The first approach identified a transposon mutant of *E. coli* O157:H7 that has lost the ability to adhere to HeLa cells and that also colonizes cattle poorly. For the second approach, a segment (*pear*) of chromosomal DNA from *E. coli* O157:H7 was cloned which, when expressed in derivatives of nonadherent *E. coli* K12, provides these strains with the ability to adhere to HeLa cells.

Our sequence analysis demonstrates that the candidate adhesin of *E. coli* O157:H7 is a homolog of the IrgA protein in *Vibrio cholerae* (Goldberg, M.B., et al., Proc. Natl. Acad. Sci. USA 88:1125-1129, 1991). This molecule is related to the TonB dependent family of outer membrane proteins in *E. coli* (Goldberg, M.B., et al., Molecular Microbiology 6(16):2407-2418, 1992). The DNA sequence analysis of the *irgA* homolog which is encoded by *pear*, and the homology to genes encoding IrgA and CIR, are described in the following Examples.

EXAMPLE 1

Identification of mutants that have lost the ability to adhere to HeLa cells.

A prototype adherent strain, *E. coli* 0157:H7 substrain 86-24, was isolated in 1986 from a patient with hemorragic colitis whose illness was traced to ground beef the patient had consumed in a fast food restaurant (23). The transposon Tn*phoA* was used to mutagenize substrain 86-24 according to the procedures in the appended citations 73 and 74, which are hereby incorporated by reference. Tn*phoA*, a Tn5 derivative, carries the alkaline phosphatase gene (*phoA*) and inserts in random locations throughout the *E. coli* genome but only in 1-3 sites per target cell. Hence, an in-frame integration of *phoA* of Tn*phoA* into regions of genes encoding an extracellular domain of a protein will result in the expression of PhoA.

PhoA-expressing Tn*phoA* mutants of *E. coli* O157:H7 were tested in the HeLa cell assay to find nonadherent mutants. This assay is described in detail in citation 8, which is hereby incorporated by reference. Bias was avoided by coding each organism, and the examining microscopist was blinded to which samples were controls and which were mutants.

EPEC strain B171 was used as the positive control for localized adherence in all experiments (58). Negative controls were nonadherent *E. coli* HB101 and/or NM554 (56). Besides the prototype adherent substrain of *E. coli* O157:H7, 6 other substrains tested adhered in a localized pattern to HeLa cells in the presence of D-mannose.

As previously reported by others (75), day to day variability in the degree of adherence of *E. coli* O157:H7 to HeLa cells was typically observed. However, three of 177 PhoA-expressing transconjugants screened for adherence to HeLa cells proved to be consistently nonadherent when tested in the coded assay (strains A5, F4, and N11). Strains A5, F4, and N11 retained all other phenotypic and genotypic characteristics of the parent strain of *E. coli* O157:H7.

Southern blot analysis determined the locations of the transposon insertions in adherent and non-adherent Tn*phoA* mutants. DNA from strains A5, F4, and N11 and from adherent mutants H8, P11, and P12 were digested with *Mlu*I, which does not cleave DNA within Tn*phoA*. Resulting fragments were separated in an agarose gel, transferred to Nytran, and probed with a fragment from the Tn5 central region of Tn*phoA*. Interestingly, the results indicated that there were two Tn*phoA* insertions in the nonadherent mutants A5 and N11, and three in F4, in apparently identical *Mlu*I bands of 23 and 16 kb length. Single integrations of Tn*phoA* are demonstrated in each of the three adherent transconjugants. Tn*phoA* integrated in the chromosome of strains A5, F4, and N11, and not in plasmid DNA.

EXAMPLE 2

Animal testing.

The nonadherent strain F4 and wild type *E. coli* O157:H7 were tested for their ability to colonize conventional Holstein calves (< 1 week old). After an initial feed of colostrum, calves were placed in individual holding pens in an isolation facility, and reared on whole cow's milk with free-choice access to water, alfalfa hay, and a high protein grain mixture. It was demonstrated at the outset that the calves were not excreting *E. coli* O157:H7 by culturing their feces on sorbitol-MacConkey agar (SMA). Four animals received either 10^8 adherent *E. coli* O157:H7 86-24 NalR or 10^9 nonadherent mutant strain F4. In dual challenge experiments, each of four calves simultaneously received 10^8 adherent *E. coli* O157:H7 86-24 and 10^9 nonadherent Tn*phoA* mutant F4. *E. coli* 0157:H7 was a spontaneously nalidixic acid resistant mutant selected on agar plates containing nalidixic acid. Tn*phoA* encodes kanamycin resistance.

The respective antibiotic resistances of these strains were exploited to identify *E. coli* O157:H7 in fecal samples by screening for shed challenge organisms on sorbitol MacConkey agar (SMA) containing nalidixic acid with or without kanamycin. Antibiotic resistant, sorbitol nonfermenting colonies were confirmed to be *E. coli* O157:H7 by their reactivity in the O157

latex particle agglutination test (Oxoid E. coli 0157 Test; Unipath Limited, Hampshire, England). The nonadherent strain was detectable for fewer days and at lower concentrations as shown in FIGURE 1, which summarizes the results of all challenges. The animals showed no ill effects which could be attributed to the *E. coli* O157:H7.

The shedding index (cfu/g of stool x number of days shed) was significantly greater for the adherent than for the non-adherent strain when analyzed by non-parametric rank sum analysis ($p=0.028$). Strain F4 grows as well as strain 86-24 NalR in fresh bovine stool and rumen contents, and in liquid broth, incubated aerobically overnight. These data suggest that the abbreviated excretion of the Tn*phoA* mutant by the challenged calves is not related to decreased viability of the mutant compared to the parent strain, even though it is difficult to simulate *in vitro* the exact conditions of the calf gastrointestinal tract. By demonstrating that calves retain the adherent strain more effectively than the nonadherent strain, these results validate the use of the HeLa cell *in vitro* adhesion assay for use in development of other reagents relevant to vaccine preparation.

EXAMPLE 3

Expression of a recombinant adhesin using chromosomal DNA from *E. coli* O157:H7.

A segment of DNA has been derived from the chromosome of *E. coli* O157:H7 strain 86-24 NalR that renders nonadherent *E. coli* NM554 adherent. To clone this segment, approximately 2000 PhoA expressing and nonexpressing Tn*phoA* mutants of *E. coli* O157:H7 86-24 NalR were screened, and one transconjugant (20D2B) was found that no longer reacted in the O157 latex particle agglutination test. This sorbitol negative mutant produced SLT II, was H7 antigen positive and β-glucuronidase negative, and possessed the same API score as the parental strain. (API score refers to a product produced by Analytab, Plainview, NY, which determines multiple bacterial growth characteristics. A score is given for each characteristic; taken in total, the score speciates bacteria. Within a species, there may be multiple scores.) However, 20D2B was highly adherent to HeLa cells. A partial *Sau*3a digest of genomic DNA of the hyperadherent strain 20D2B was ligated into plasmid Supercos (pSC) (20,77), packaged, and used to transduce nonadherent laboratory strain NM554. This experiment yielded 2200 transductants with an average of 40 kb of DNA inserted into the *Bam*HI site of pSC.

The 2200 cosmid clones were screened for adherence to HeLa cells, and two adherent clones were identified and designated pSC(A-G6) and pSC(T-H12). *E. coli* NM554 containing pSC(A-G6) and pSC(T-H12)) adhered to HeLa cells in a diffuse rather than localized pattern although nascent clusters were sometimes seen. Southern blotting demonstrated that: (a) the A-G6 and T-H12 determinants overlap by approximately 15 kb; (b) these inserts are derived from *E. coli* O157:H7 chromosomal DNA; (c) the inserts do not

encode *eae*, *bfp* (which encodes the bundle forming pilus adhesin of EPEC), or SLT II; and (d) the overlap region is conserved in each of 9 *E. coli* O157:H7 tested, but not in *E. coli* HB101, DH5α, or EPEC B171.

As shown in FIGURE 2, a deletion mutant of pSC(A-G6), designated "pSC(overlap)" (ATCC No. 69648), retains the overlapping segment between pSC(A-G6) and pSC(T-H12). Interestingly, nonadherent *E. coli* HB101 transformed with pSC(overlap) display diffuse adherence to Madin-Darby bovine kidney cells (MDBK), and, in a preliminary experiment, localized adherence to HeLa cells. An 8 kb subclone of pSC(overlap), designated "p*ear*", restores adherence to non-adherent strain A5. (p*ear*, and the *irgA* homologous subclone described below, display diffuse adherence to HeLa cells.)

The data summarized above suggest that: (1) an identifiable adhesin from *E. coli* O157:H7 expressed in *E. coli* HB101 (p*ear*) enables *E. coli* O157:H7 to adhere to epithelial cells of human (HeLa) and bovine (MDBK) origin *in vitro*; and (2) this adhesin is the same molecule which permits *E. coli* O157:H7 to remain in the gastrointestinal tracts of bovines. Further identification and characterization of the subject recombinant *E. coli* 0157:H7 adhesin is described below.

EXAMPLE 4

Identification of the genes on the adherence conferring plasmid (p*ear*).

pSC(overlap) itself consists of 4 kb of pSC DNA and approximately 15 kb of *E. coli* O157:H7 DNA. p*ear* consists of 8 kb of chromosomal DNA plus the SK+ vector (Stratagene). To identify the adhesin expressed by p*ear*, the entire fragment was sequenced and open reading frames were determined. The results are described below.

The appended SEQ ID NO:1 shows the 8,041 base pair nucleotide sequence of p*ear*. Almost all of the sequence has been confirmed. Ambiguous DNA (in regions not encoding the candidate adhesin) is noted by N in the appended sequence. The p*ear* insert contained three open reading frames (ORFs) of sufficient length to encode potential virulence or adherence factors. Two of these are homologous to genes necessary for resistance to tellurite (Jobling, MG, et al., Gene 66:245-258, 1988). These *terE* and *terD* homologs are shown in SEQ ID NO:2 and SEQ ID NO:3, corresponding respectively to nucleotides 7024-6449 and 7670-7092 of SEQ ID NO:1. The other ORF is homologous to a gene encoding a homolog of IrgA (Goldberg, MB, et al., Molecular Microbiology 6:2407-2418, 1992). This *irgA* homolog is shown in SEQ ID NO:4, which corresponds to nucleotides 3036-5126 of SEQ ID NO:1. IrgA is an outer membrane protein of *V. cholerae*, and is believed to be important for colonization of mice in an experimental system (Goldberg, MB, et al, Infection and Immunity, 58:55-60, 1990). The *E. coli* O157:H7 adhesin (SEQ ID NO:4) is also homologous to the *E. coli* colicin I receptor (CIR) (Griggs, D.W., et al., J. Bacteriol.

168:5343-5352, 1987). The amino acid homologies of the candidate adhesin to IrgA and to CIR are demonstrated by comparing SEQ ID NO:5 and SEQ ID NO:6, and SEQ ID NO:7 and SEQ ID NO:8, respectively.

EXAMPLE 5

Mutations in the *irgA* homolog, as cloned into an expression vector, lead to loss of adherence. Transposon (Tn*phoA*) insertions in the *irgA* homolog of *E. coli* O157:H7 ablate adherence of laboratory strains of *E. coli* transformed with a plasmid vector into which an adherence conferring region has been inserted. We cannot state with certainty the exact site of the two Tn*phoA* insertions which ablated adherence, but the regions of the insertions are between nucleotides 3271-3310 and 3801-3840 of SEQ ID NO:1.

EXAMPLE 6

A product of a single gene (i.e., the *irgA* homolog) confers adherence to nonadherent *E. coli*. We first performed PCR using as primers the sequences 5'GGGGATCCAATTCTGGCATGCCGAGGCAGTCG3' (SEQ ID NO:9), corresponding to nucleotides 2895-2914 of SEQ ID NO:1) and 3'GGACCGCCTTGTCACCGTTGCTCTTAGATCTGG5' (SEQ ID NO:10, corresponding to nucleotides 5176-5196 of SEQ ID NO:1) from which DNA on p*ear* was amplified. These sequences were cloned into the *Bam*HI and *Xba*I sites of pSK+. We also amplified the same gene using as template DNA from *E. coli* O157:H7. In this latter case,. the primers used were 5'GGAAGGATCCCCGAACACGCCATACGGATAGCTG3' (SEQ ID NO:11, corresponding to nucleotides 2867-2890 of SEQ ID NO:1) and 3'GCAACGGTGACGTTGAGGACCGCCAGATCTAAAGG5' (SEQ ID NO:12, corresponding to nucleotides 5159-5183 of SEQ ID NO:1). This latter PCR product was also cloned into pSK+, using the same *Bam*HI and *Xba*I sites. In both cases, multiple laboratory strains of nonadherent *E. coli* were rendered adherent to HeLa cells by these cloned single genes.

EXAMPLE 7

<u>The adherence of Δ-*ear* mutants to HeLa cells is diminished.</u>

Strain F12 of *E. coli* O157:H7 is a hyperadherent mutant that has been mutated by Tn*phoA* such that the O157 antigen is no longer expressed (Bilge, S.S., et al., Abstract B-7, American Society of Microbiology, 21-25 May 1995). F12 is probably hyperadherent because the lack of expression of the O157 antigen enables the adhesin to be more completely exposed on the bacterial cell surface.

We deleted the entire 8041 base pair *Kpn*I-*Kpn*I region (SEQ ID NO:1) of p*ear* from strain F12 as follows. We cloned p*ear* into a suicide vector, pCVD442 (Donnenberg, MS, et al., Infection and Immunity, 59:4310-7, 1991), which was then mated into strain F12 using

E. coli SM10 lambda pir as a donor. Sucrose resistant, ampicillin sensitive strains were analyzed to find mutants with the SEQ ID NO:1 region deleted.

When the *ear* is deleted from strain F12, the organism was observed to be nonadherent or severely adherence deficient (1 cluster of microcolonies per 1-3 high powered fields) by an observer blinded to the identity of the F12 and the *ear* deletion mutant. (In comparison, the parent strain F12 displayed much higher levels of adherence to HeLa cells, approximately 0.5-1 cluster per cell.) This striking adherence deficiency could be complemented by the cloned genes of *irgA* from either the plasmid or the chromosome. Hence, this loss of adherence from the deletion of the *ear* is not caused by a polar effect of the deletion.

In summary, our data demonstrates that the PCR product of a single allele, an *irgA* homolog in *E. coli* O157:H7, confers an adherent phenotype when cloned into an appropriate vector and transformed into laboratory strains of *E. coli*. Tested strains include: *E. coli* NM554 (Raleigh, EA, et al, Nucleic Acids Research, 16:1563-75, 1988); *E. coli* HB101; and *E. coli ORN172* (Woodall, LD, et al., Journal of Bacteriology 175: 2770-8, 1993), which is an *E. coli* K12 strain from which genes encoding type I pili have been deleted. Our deletion mutation data confirm that the epithelial adherence region (*ear*) encodes an *E. coli* O157:H7 adhesin. Sequence data suggest that this adhesin is a homolog of IrgA of *V. cholerae*.

We have also performed Tn*phoA* mutagenesis of *E. coli* O157:H7, and identified three nonadherent mutants (strains A5, F4, and N11), each of which sustained a Tn*phoA* insertion in the same allele (SEQ ID NO:9). One of these strains, strain F4, was deficient in its ability to colonize in calves in an oral challenge experiment performed at the Washington State University in Pullman, Washington. Sequence analysis suggests that the Tn*phoA* insertion in the same allele among the three nonadherent mutants may have taken place in the midst of a cluster of genes, at least one of which has homology to pro-secretory proteins in *Yersinia enterocolitica* (YscJ) (Michiels, T., et al., Journal of Bacteriology 173:4994-5009, 1991), *Rhizobium fredii* (Nolt) (Meinhardt, L.W., et al., Molecular Microbiology 6:2407-2418, 1992), and *Xanthomonas compestris* (HrpB) (Fenselan, S., et al., Molecular Plant-Microbe Interactions 5:390-396, 1992), and it is possible that the secretion of the *E. coli* O157:H7 adhesin is controlled by this secretory mechanism.

EXAMPLE 8

Construct recombinants and deletion mutants for bovine challenge experiments.

Our data suggest that adherence to HeLa cells by *E. coli* O157:H7 correlates with optimal colonization of calves with this organism. However, the *E. coli* O157:H7 adhesin is not closely linked to a locus mutagenized by Tn*phoA* in nonadherent strain F4. To exclude the possibility that separate *in vitro* and *in vivo* adherence mechanisms exist for *E. coli* O157:H7, we may (1) colonize calves with a laboratory *E. coli* expressing the recombinant

O157:H7 adhesin; (2) immunize animals with a recombinant adhesin and ascertain if these animals are protected from challenge with *E. coli* O157:H7; or (3) create an isogeneic deletion mutant of *E. coli* O157:H7 and determine if this strain has lost its ability to adhere to HeLa cells and to colonize calves.

EXAMPLE 9

Study adhesin deletion mutants of *E. coli* O157:H7 in calves.

We shall determine if adhesin-deletion mutants of *E. coli* O157:H7 lose the ability to colonize challenged animals in the calf model. We have demonstrated that wild type *E. coli* O157:H7 colonizes the gastrointestinal tracts of calves longer and at higher concentrations than does a nonadherent Tn*phoA* mutant. Ideally, we would challenge calves with a laboratory strain of *E. coli* expressing the recombinant adhesin of *ear*, and follow excretion of this organism. However, it is very unlikely that a recombinant host strain will survive in the gastrointestinal tract of animals. For example, plasmids of enterotoxigenic *E. coli* encoding K88 (F4) or K99 (F5) pilus adhesins did not confer the ability to colonize pig small intestine when transferred to laboratory *E. coli* strain K12, despite the clearly demonstrated role of these adhesins to colonizations of the small intestines in the wild type strains (48). Therefore, we use specific deletion mutants to authenticate the role of the *in vitro* adhesin in the calf model.

To validate the role of the *ear* adhesin in bovine colonization, we created an isogeneic, in-frame deletion mutant of almost all of the *irgA* homolog in *E. coli* O157:H7. To do this, we cloned regions (ca. 0.45 kb) flanking the gene encoding the IrgA homolog adhesin, up to and including nucleotide 3110 at the 5′ region, and nucleotide 5051 at the 3′ region of the gene encoding the IrgA homolog adhesin, such that the 1941 nucleotides between nucleotides 3110 and 5051, inclusive, are deleted from this construct, corresponding to 647 amino acids. These truncated sequences of the IrgA homolog adhesin were then inserted into pCVD442, and deletion mutants of strain 86-24 NalR were identified by virtue of their ampicillin sensitivity and sucrose resistance. The deletion was confirmed in two mutants by Southern blotting using as probes the internal (i.e., deleted) portion of *irgA* homologs and DNA flanking the deleted sequences. The deletion mutant is termed 86-24 (Δ *irgA*). The parent strain (86-24 NalR) and a ten-fold excess of deletion mutant 86-24 (Δ *irgA*) will be fed simultaneously to cattle, and fecal excretion will be monitored by culturing stool on SMA plates containing nalidixic acid. A representative number of strains will be collected at multiple time points after challenge, until clearance of the challenge strains, or until 60 days post challenge, whichever comes first. These strains will be individually tested for the presence of the *irgA* deleted sequences using the polymerase chain reaction, to determine which of the strains colonizes the cattle.

In a representative experiment, animals are challenged with dual inocula (adherent and excess nonadherent deletion mutant of *E. coli* O157:H7). Based on previous challenges, we can anticipate that at least 6 calves will excrete *E. coli* O157:H7 for at least 21 days after inoculation.

Study the humoral and secretory immune response of the calves challenged with *E. coli* O157:H7: The immune response of bovines to enteric colonization with *E. coli* O157:H7 is unknown. The calf challenge experiments described above provide a model to study this response. In particular, the humoral and secretory immune response to various *E. coli* O157:H7 components, including but not limited to the recombinant adhesin, can be quantified to determine if the development of a response is correlated to the clearance of this organism from the gastrointestinal tract.

To study the immune response of the challenged bovines to *E. coli* O157:H7, serum and saliva are obtained from each calf prior to challenge and at each sampling time, and stored at -70°C. Intestinal secretions are saved at the end of the experiment, when animals are necropsied.

The humoral and secretory immune responses of the challenged bovines to *E. coli* O157:H7 proteins are studied by protein immunoblot. Outer membranes, as well as whole cell preparations, of *E. coli* O157:H7 86-24, *E. coli* NM554 pSC(A-G6), *E. coli* NM554 pSC(T-H12), and *E. coli* HB101 (*pear* or a derivative) are electrophoresed in adjacent lanes in 10% SDS-PAGE gels and transferred electrophoretically to nitrocellulose membranes (32). Outer membranes from nonadherent *E. coli* are included as controls. The membranes are probed with calf serum and intestinal secretions, and adherent primary antibodies are detected by murine monoclonal antibodies to bovine IgG1, IgG2, IgM, and IgA, and goat antimouse IgG coupled to horseradish peroxidase. In control blots, the monoclonal antibodies are replaced with isotype-matched monoclonal antibodies to irrelevant antigens.

Protein immunoblot studies are extended to the recombinant adhesin produced above using pGEX (67) as well as to other strains of *E. coli* O157:H7. Antibodies from calves challenged in this section are also used in an attempt to inhibit *in vitro* adherence of *E. coli* O157:H7 strain 86-24 and the recombinant adherent strains to epithelial cells. Finally, the recombinant adhesin is an appropriate target molecule for an enzyme-linked immunosorbent assay to quantify the magnitude of the antibody response to the adhesin in cattle challenged with *E. coli* O157:H7.

EXAMPLE 10

Immunological characterization of a recombinant *E. coli* 0157:H7 adhesin.

Standard immunochemical techniques are used to determine if the cloned adhesin (SEQ ID NO:4) is the same as the adhesin used by *E. coli* 0157:H7 to adhere to HeLa cells.

To achieve this objective, outer membrane proteins are prepared from laboratory strain(s) expressing the recombinant adhesin (SEQ ID NO:4). These proteins are analyzed on SDS-polyacrylamide gels, and used to immunize rabbits at three one-month intervals and Holstein cows at 30 and 60 days prepartum. Rabbits are also be immunized with killed whole bacterial cell preparations. A recombinant *E. coli* NM554, which does not adhere to HeLa cells, is used as a negative control immunogen.

Decomplemented serum from immunized rabbits and cows, and whey from the milk of immunized cows, is first absorbed against *E. coli* NM554 to deplete antibodies to outer membrane proteins which are irrelevant to adherence. The absorbed sera and milk are tested against *E. coli* 0157:H7 strain 86-24, as well as additional strains, using immunofluorescence and Western blotting to determine if antibody to the recombinant antigen has specific affinity for extracellular antigens or outer membrane proteins, and to determine if this antigen is conserved. Negative control antigens consist of other *E. coli*, including diarrheagenic strains, as well as other Shiga like toxin-producing *E. coli* which do not belong to serotype 0157:H7. Additional negative antibody controls include preimmune serum, and serum and milk obtained from animals immunized with nonadherent recombinant *E. coli* NM554.

The function of anti-adhesin antibodies is assessed using the *in vitro* HeLa cell assay, adapted to quantify the numbers of bacteria adherent to the target cells. *E. coli* 0157:H7, as well as adherent and nonadherent recombinant *E. coli* NM554, are incubated with immune or control sera or milk antibodies before addition to the HeLa cell culture. Antibodies remain in the adherence assay medium. Additionally used in these assays are antibodies in serum and saliva from animals challenged with oral *E. coli* 0157:H7. After the appropriate incubation period, the number of bacteria adherent per cell is enumerated in multiple fields consisting of several hundred eukaryotic cells. The microscopist is blinded to the identity of the strains and antibodies.

The anti-adhesin antibodies raised and selected as described above are also useful for the diagnostic identification of *E. coli* O157:H7. So too are *E. coli* O157:H7 nucleotide sequences within or flanking the *irgA* homolog having the requisite specificity and sensitivity for diagnosing the presence of strain O157:H7 in feed animals, food, and humans, as determined by screening a panel of closely related bacterial strains for specificity, and a panel of *E. coli* O157:H7 for sensitivity.

EXAMPLE 11

Immunoprophylactic vaccines.

Bacterial adhesins have been used as immunogens to prevent colonization of mucosal surfaces and/or disease in multiple food and laboratory animals. An immunoprophylactic approach to the problem of *E. coli* O157:H7 carriage using a purified recombinant adhesin

(*ear* or subclone thereof) as a vaccine is considered to be an efficient method to improve the microbiologic safety of food of bovine origin. Preparation and administration protocols for such vaccines are known in the art. See, e.g., U.S. Patents No. 5,286,484; No. 5,208,024; No. 5,137,721; No. 5,079,165; No. 5,066,596; No. 4,795,803; No. 4,736,017; No. 4,702,911; No. 4,725,435; No. 4,454,116; and No. 3,975,517, which are incorporated by reference herein.

It is well known in the art that vaccines administered to cattle can confer immunity to bacteria that colonize the gut. For example, U.S. No 3,975,517, describes operable methods for vaccinating cattle. Other published methods suitable for immunizing animals include: Acres et al., 25 *Infection and Immunity 121*, 1979; Linggood et al., U.S. 4,971,794, 1990; Sadowski and Prairie, U.S. No. 4,652,448, 1987; Suchaux et al., 69 *Infection and Immunity* 2828, 1992; Francis and Willgohs, 52 *Am. J. Vet. Res* 1051, 1991; Ikemori et al., 53 *Am. J. Vet. Res.* 2005, 1992; Isaacson et al., 29 *Infection and Immunity* 824, 1980; Morgan et al., 22 *Infection and Immunity* 771, 1978; Morris et al., 13 *J. Med. Microbiol.* 265, 1980; Runnels et al., 55 *Infection and Immunity* 555, 1987; Sojka et al., 11 *J. Med. Microbiol.* 493, 1978; and Yokoyama et al., 60 *Infection and Immunity* 998, 1992. Vaccines can be prepared and administered by a variety of routes, many of which are found in Harlow and Lane, *Antibodies*, 1988, which is hereby incorporated by reference.

Immunization with the *ear*-encoded IrgA homolog adhesin to result in decreased excretion of *E. coli* O157H7 requires generation of immune responses with effectors in the gastrointestinal tract of adult and sub-adult cattle, the age groups which are presented for slaughter. In this age group, the presence of a functional rumen presents a major barrier to oral vaccination in a form designed to survive passage through the abomasum (the 'true' stomach), as the very large and metabolically active rumenal compartment is thought to block passage of most intact antigens to the abomasum.

Several strategies can result in immune responses in the intestine in adult ruminants. These include: the intramammary route mentioned below; parenteral (intramuscular) immunization, because the resulting antibodies are cleared from circulation into the intestine; intranasal immunization (as antigen presentation by the pharyngeal tonsils is quite effective); and the use of highly stimulating adjuvants such as ISCOMs (immune stimulating complexes such as liposome-antigen combinations) or cholera toxin b-subunit conjugates.

Types Of Vaccines: Purified antigen (prepared using standard recombinant DNA methods) or whole-cell vaccines can be used to stimulate an immune response in pregnant cows, resulting in the presence of protective antibodies in the colustrum and milk produced after parturition. The milk or colustrum can be stored for later administration to newborn calves. Alternatively, protein produced from the *irgA* homolog can be used to raise monoclonal antibodies which then can be used directly to confer passive immunity to newborn

calves. Methods for preparing monoclonal antibodies are well-known in the art, and can be found, for example, in Harlow and Lane, *Antibodies*, Cold Spring Harbor Laboratory, 1988.

Vaccines Containing Purified Antigen: Purified IrgA homolog adhesin is formulated in sesame or peanut oil, sterile aqueous solution, or other isotonic solution suitable for injection. Stabilizers such as sorbitol or gelatin may be added. The immunogenicity can be enhanced by including various immunoadjuvants in the pharmaceutical preparation. Immunoadjuvants useful for this purpose include the water-in-oil emulsion of Freund, light mineral oil (Bayol), the commercially available emulsifiers Arlacel A and Aracel C, or monophosphoryl lipid A.

Cows are inoculated subcutaneously above the left shoulder twice, approximately 3 and 6 weeks before the onset of calving.

Whole Cell Vaccines: A bacterial strain suitable for use as a whole-cell vaccine is a nonpathogenic (to animals and humans) bacterium expressing the homolog adhesin (SEQ ID NO:4). This strain can be administered orally as a live vaccine, or can be used as a killed vaccine. Oral vaccination is an efficient means for stimulating production of secretory IgA in mucosal tissues such as the internal lining of the intestinal tract. Using a killed vaccine is the safest choice, as it obviates the risk that the non-pathogenic vaccine strain will acquire virulence genes by genetic transfer from other bacteria that may be present in the gut.

Where the goal is to acquire materials suitable for providing passive immunization, another advantageous route of administration is to inject the vaccine preparation directly into the teat canal of pregnant or nursing cows. Whether administration is oral or via injection, the vaccinated animal will produce protective antibodies in her milk that will passively immunize recipient calves.

Preparation Of Formalin-Treated Vaccine: One means of vaccine preparation incorporates some of the advantages of both live and killed whole-cell vaccines. For this procedure, a bacterial strain expressing the antigenic protein is subjected to a controlled formalin treatment. A broth culture of the vaccine strain is grown to a concentration of about 10^8 to 10^9, then incubated under aerobic conditions for 10 to 15 hours in the presence of about 0.04% (vol/vol) formalin (0.016% wt/vol formaldehyde). Alternatively, the vaccine strain is grown on solid media, and the cells scraped off and suspended in broth medium for the formalin treatment. Formalin-treated bacteria remain viable, but optimally the proportion of colony-forming units is reduced by about 1000-fold compared with bacteria not exposed to formalin. For each batch of vaccine, plate counts are performed to ensure that the requisite proportion of bacteria have survived the formalin treatment. If necessary, the time of exposure to formalin and the percentage of formalin added to the broth are adjusted so that the plate counts of post-treatment cultures are 1000-fold reduced compared with controls. The vaccine contains the entire broth culture constituents including metabolic waste products and extracellular proteins.

Virulence of vaccine strains are tested by oral inoculation of 3 to 4 day old suckling mice with either virulent *E. coli* O157:H7 or the formalin-treated vaccine strain. Each mouse is administered about 10^6 to 10^8 organisms in about 0.15 ml. Mice are orally inoculated with serial dilutions of the bacteria, and survival is determined at 40 hours.

Administration Of Vaccine: For oral administration, the formalin-treated vaccine strain is packaged in enteric coated capsules that dissolve only after passing through the stomach. To achieve prophylaxis with monoclonal antibodies, the antibody can be mixed with colostrum, or any pharmaceutical carrier suitable for oral administration.

Injection: Four to eight ml doses of the vaccine are administered intramuscularly or directly into the lactiferous ducts via the teat canals of two year old cows. For lactiferous duct administration, the cows are given a series of 4 to 6 ml doses at 42, 32, 22 and 12 days prior to parturition. For intramuscular vaccination, 4 ml of the formalinized culture is injected in the side of the neck 28 days before parturition, followed by 6 ml 14 days later, and 8 ml at 5 days before the expected calving day. About 1 liter of milk is collected from vaccinated and control cows seven days after parturition. Whey is prepared by centrifuging milk for 44,000 g for two hours, then collecting the supernatant. Whey is stored either in 2 ml aliquots at 4°C, stored frozen at -30°C, or stored lyophilized.

To monitor antibody production in response to the vaccine, 1 l of milk is taken from vaccinated pregnant cows on 2,3,4,7 and 14 days after calving. Whey is prepared from the milk by centrifugation and stored as described above.

Challenging Passively-Immunized Calves: To demonstrate the protective effects of whey from vaccinated teats, newborn calves separated from their mothers are fed one aliquot of post-immune whey every 8 hours for 72 hours. Control calves are fed whey from non-vaccinated cows. Six hours after being fed the first dose of whey, the calves are infected orally with broth culture containing about 10^9-10^{10} colony forming units virulent *E. coli* O157:H7 per ml. Each day after being exposed to the O157:H7, stool samples from infected calves are cultured to detect the presence or absence of *E. coli* O157:H7. Three calves from each group are sacrificed after 10 days, and their intestines examined for the presence or absence of *E. coli* O157:H7.

Mucosal Immunization: An additional route of vaccination exploits the mucosal immune system, which provides a defense against colonization of surfaces lining the gastrointestinal and respiratory systems, and relies on the induction of IgA antibodies secreted from these organs into the lumens where organisms colonize, and viable bacterial vectors which are in themselves not harmful to animals or humans, but which could be genetically engineered to express the *E. coli* O157:H7 adhesin. For a review of mucosal immunization, see: McGhee, J.R., et al., Seminars in Hematology 30 (4 suppl 4):3-12, discussion 13-15,

1993. For example, an attenuated or harmless commensal organism, including but not limited to a member of the *Enterobacteriaceae*, is transformed with the *E. coli* O157:H7 recombinant adhesin. This construct is administered to animals orally, or via intranasal inoculation. The efficacy of this approach is measured by assaying for specific IgA to the antigen, or by challenging vaccinated and control animals with *E. coli* O157:H7, and following excretion of this pathogen.

Immunization With Recombinant Bacterial Antigen Produced in Transgenic Plants: An alternative approach to the prevention of the carriage of *E. coli* O157:H7 by food animals, and possibly by humans, is to introduce the *E. coli* O157:H7 adhesin expressed in transgenic plants. For example, Haq et al. introduced sequences encoding *E. coli* heat-labile enterotoxin (LT-B) on an expression vector into *Agrobacterium tumifaciens*. This host bacterial strain was used to transfer the sequences encoding LT-B into tobacco (*Nicotiana tabacum* cv. Samsun) and potato (*Solanum tubrosum* variety "Frito-Lay 1607") plants, which then expressed variable amounts of these toxin antigens, as determined by immunoassay. These bacterial proteins were fed to mice, either in the form of extract of tobacco plants which expressed LT-B, or as potatoes expressing LT-B (i.e., no extract). In both cases, serum IgG and mucosal IgA responses to LT-B were elicited.

Thus, as an analogous approach in the immunization of cattle to prevent carriage of *E. coli* O157:H7, transgenic plants expressing the *E. coli* O157:H7 adhesin are fed to food animals. The response to the antigen is measured by assaying for circulating IgG and mucosal IgA antibodies specific for the *E. coli* O157:H7 adhesin. These animals could then be challenged with *E. coli* O157:H7, and excretion followed, with appropriate control animals also challenged, to determine if an immunoprotective response has been elicited. The determination of immunoprotection would be analogous to the challenge with the *irgA* homolog deletion mutants, but in this case, animals immunized with transgenic plants expressing the IrgA homolog adhesin, or a control (i.e., irrelevant) antigen will be challenged with nalidixic acid resistant *E. coli* O157:H7 The excretion of challenge organisms will be followed by plating stool on SMA with nalidixic acid, and the duration and level of shedding of *E. coli* O157:H7 in the two groups will be compared.

It is also contemplated that immunizing transgenic plants may be used to protect humans from infection with *E. coli* O157:H7 and related pathogens expressing the *E. coli* O157:H7 adhesin, or homologs of this adhesin. In the case of humans, IgA and IgG can be assayed to detect a humoral response to this protein, but challenge experiments cannot be performed because of the exceptional pathogenicity of *E. coli* O157:H7.

Use of a Recombinant Adhesin as a Competitive Inhibitor of *E. coli* O157:H7 in the Gastrointestinal Tracts of Food Animals and of Humans: Large quantities of recombinant adhesin expressed by plants and consumed in the diets of animals and humans may compete

with wild type *E. coli* O157:H7 and related pathogenic bacteria for binding sites on the enterocyte surface. For example, animals are put on such feed in the several days or weeks prior to shipment for slaughter, provided controlled challenge experiments show that the recombinant adhesin, expressed in plants which are then fed to the animals, promotes clearance of *E. coli* O157:H7 from the gastrointestinal tract of such animals, thereby reducing the load of this pathogen that enters the production line in abattoirs. In a similar approach, children with the early stage of gastrointestinal infection with *E. coli* O157:H7 are administered these recombinant competitive inhibitors, including but not limited to the recombinant adhesin expressed on plants, to promote clearance of the organism, thereby ameliorating infection, or preventing the development of hemolytic uremic syndrome (HUS). Also, contacts of children with *E. coli* O157:H7 infection, who have yet to display symptoms, are fed recombinant adhesin, with the intent that colonization in such potential secondary cases is prevented.

The disclosed adhesin may also be used to prevent or ameliorate human infection with *E. coli* O157:H7, as well as with bacteria which use a homolog of this adhesin, with shared active sites or epitopes. As discussed above, the *E. coli* O157:H7 adhesin demonstrates striking homology to IrgA. There are domains of high degree of homology (conserved regions) interspersed with relatively nonconserved regions (variable regions). The possibility exists, therefore, that the IrgA homolog in *E. coli* O157:H7 can be used as an immunogen (vaccine) against *V. cholerae* infections. An additional possibility is that the IrgA homolog in *E. coli* O157:H7 may be a useful vaccine against the carriage of other pathogenic *Enterobacteriaceae* by food animals, such as diarrheagenic *E. coli* that do not belong to serogroup O157:H7.

The IrgA homologue might also be useful in the induction of protective immunity, or as a competitive inhibitor of colonization, in several additional hosts and in several additional infections. These include (1) a vaccine to prevent infections in humans caused by *E. coli* O157:7H, other diarrheagenic *E. coli*, and *V. cholerae*; (2) a vaccine to prevent carriage by cattle, and other animals destined for human food, of diarrheagenic *E. coli*; (3) a competitive inhibitor of colonizing diarrheagenic *E. coli* in food animals. To test the efficacy of these approaches, multiple experiments are contemplated. These include (1) the administration of immune globulin (passive immunity) from donors immunized with the IrgA homolog adhesin to high risk patients (i.e., children in contact with a primary case of *E. coli* O157:H7 infection), in a placebo-controlled fashion, to determine the differences in attack rates between the two groups; (2) a vaccine formulation could be administered to human volunteers subsequently challenged with *V. cholerae*, or to patients residing in areas of high risk for infection; (3) challenge of immunized food animals with related diarrheagenic *E. coli* (i.e., not *E. coli* O157:H7) and determine if colonization can be established. The possibility also exists

that animals infected with pathogenic *E. coli* sharing an IrgA-mediated adherence mechanism cause illness in the animal industry. The IrgA homolog adhesin may, in these cases, be used to diagnose, prevent, or treat the infection. In fact, we have identified *pear* sequences in RDEC-1 and some *E. coli* strains from calves with diarrhea.

Glossary

Glossary of abbreviations, strains, plasmids, and genes relevant to this disclosure:

EPEC	enteropathogenic *E. coli*
MDBK	Madin-Darby bovine kidney cells
PhoA	alkaline phosphatase
PBS	phosphate-buffered saline
pSC	plasmid Supercos
PCR	polymerase chain reaction
SLT	Shiga-like toxins
SLT I	Shiga-like toxin I
SLT II	Shiga-like toxin II
SMA	sorbitol-MacConkey agar
XP	5-bromo-4-chloro-3-indolylphosphate

E. coli O157:H7 86-24: An SLT-II producing isolate which caused a large restaurant-associated outbreak of hemorrhagic colitis which included two deaths (23,71).

E. coli O157:H7 NalR: Nalidixic acid resistant adherent mutant of *E. coli* O157:H7 86-24 used in Tn*phoA* mutagenesis. This strain adheres in a localized pattern to HeLa cells.

E. coli O157:H7 Strains A5, F4, N11: Tn*phoA* mutants of *E. coli* O157:H7 86-24 NalR which express PhoA, and are non-adherent, kanamycin resistant, and ampicillin sensitive.

E. coli O157:H7 Strains H8, P11, P12: Tn*phoA* mutants of *E. coli* O157:H7 86-24 NalR which are adherent, kanamycin resistant, and ampicillin sensitive.

E. coli 20D2B: a Tn*phoA* mutant of *E. coli* O157:H7 NalR which is hyperadherent. This strain does not express the O157 antigen, but retains all other characteristics of the parent strain (H7 antigen positive, sorbitol nonfermenting, SLT II positive, same API score).

E. coli NM554: A laboratory strain of *E. coli* used for cosmid cloning.

Plasmids used:

pSC(A-G6) and pSC(T-H12): overlapping cosmid clones containing 30 kb of *E. coli* O157:H7 chromosomal DNA expressed in *E. coli* NM554 which confers D-mannose-resistant

adherence on this *E. coli* in a mostly diffuse pattern.

pSC(overlap): deletion mutant of pSC(A-G6) which retains 15 kb overlap region, and confers adherence to *E. coli* HB101. pSC(overlap) was deposited on June 24, 1994, under accession number 69648 at the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852, U.S.A.

pGEX: A vector which allows the expression of the IrgA homolog adhesin and a molecule which can act as a ligand for an affinity purification step (67).

pear: adherence-conferring 7 kb subclone of pSC(overlap).

irgA: Iron regulated gene A, from *V. cholerae*. This has regions of homology to the adhesin of *E. coli* O157:H7 that was cloned and is described in this application.

CIR: *E. coli* colicin I receptor. *E. coli* outer membrane protein which also has regions of considerable homology to the described *E. coli* O157:H7 adhesin.

Strain F12: another O157 Tn*phoA* mutant which has lost the ability to express the O157 antigen by virtue of having sustained a Tn*phoA* insertion in the *rfb* locus. This strain is hyperadherent. *ear* deletions from F12 are considerably less adherent than the parent strain.

Strain 86-24 NalR (Δ *irgA*): a deletion mutant of *E. coli* O157:H7 strain 86-24 NalR, which is used in calf challenge experiments.

ear (epithelial adherence region): subclone of 15 kb pSC(overlap) region between pSC(A-G6) and pSC(T-H12), representing approximately 7 kb of cloned *E. coil* O157:H7 chromosomal DNA which confers the adherence phenotype to nonadherent laboratory *E. coli*.

Citations

1. Acres, S.D., et al., *Infect Immun* 1979; 25:121-126.
2. Baga, M., et al., *Escherichia coli. EMBO J* 1985; 4:3887-3893.
3. Bakker, D., et al., *Escherichia coli. Mol Microbiol* 1991; 5:875-886.
4. Beebakhee, G., et al., *FEMS Microbiol Lett* 1992; 91:63-68.
5. Bell, B, personal communication.
6. Belongia, E.A., et al., *JAMA* 1993; 269:883-888.
7. Bilge, S.S., et al., *J Bacteriol* 1989; 171:4281-4289.
8. Bokete, T.N., et al., *Gastroenterology* 1993; 105:1724-1731.
9. Borczyk, A.A., et al., *Lancet* 1987; 1:98.
10. Caprioli, A, et al., *J Infect Dis* 1992; 166:154-158.
11. Caprioli, A, et al., *J Infect Dis* 1994; 169:208-211.
12. Cordovez A., et al., J Clin Microbiol 1992; 30:2153-2157.
13. Cravioto, A., et al., *Lancet* 1991; 337:262-264.
14. Donnenberg, M.S., et al., *Infect Immun* 1990; 58:1565-1571.
15. Donnenberg, M.S., et al., *Infect Immun* 1991; 59:4310-4317.
16. Donnenberg, M.S., et al., *Infect Immun* 1992; 60:3953-3961.

17. Donnenberg, M.S., et al., *J Clin Invest* 1993; 92:1418-1424.
18. Duchet-Suchaux, M., et al., *Infect Immun* 1992; 60:2828-2834.
19. Enteric Diseases Branch, CDC, *Morbid Mortal Wkly Rep* 1993; 42:85-86.
20. Evans, G.A., K. et al., *Gene 1*989; 79:9-20.
21. Francis, D.H., et al., *Am J Vet Res* 1991; 52:1051-1055.
22. Fratamico, PM, et al., *J Med Microbiol* 39:371-381, 1993.
23. Griffin, P.M., et al., *Ann Intern Med* 1988; 109:705-712.
24. Griffin, P.M., et al., *Epidemiol Rev 1*991; 13:60-98.
25. Hancock, D.D., et al., *Epidemiology and Infection* 113:119-207, 1994.
26. Hancock, D.D., et al., National Prevalence Study for *Escherichia coli* O157:H7 in United States Dairy Calves. Submitted.
27. Hancock, R.E.W., et al., *J Bacteriol* 1978; 136:381-90.
28. Henikoff, S., *Gene* 1984; 28:351-359.
29. Ikemori, Y., et al., *Am J Vet Res* 1992; 53:2005-2008.
30. Isaacson, R.E., et al., *Infect Immun 1*980; 29:824-826.
31. Jacobs, A.A.C., et al., *J Bacteriol* 1987; 169:735-741.
32. Johnstone, A., et al., *Immunochemistry in Practice*, 2nd Ed. Blackwell Scientific Publications, Oxford, 1987, pp 190-196.
33. Junkins, A., et al. *Curr Microbiol* 1989; 19:21-27.
34. Karch, H., et al., *Infect Immun* 1987; 55:455-461.
35. Karmali M.A., et al., *J Infect Dis* 1985;151:775-782.
36. Kimura, A., et al., *Infect Immun* 1990; 58:7-16.
37. Krogfelt, K.A., *Rev Infect Dis 1*991; 13:721-735.
38. LeSaux, N., et al., *J Infect Dis* 1993; 176:500-502.
39. Lindberg, F., et al., *Nature 1*987; 325:84-87.
40. Lopez E.L., et al., *J Infect Dis* 1989;160:469-475.
41. Louie, M., et al., *Infect Immun* 61:4085-4092, 1993.
42. MacDonald K.L., et al., *JAMA 1*988;259:3567-3570.
43. Marshall, B., et al., *Proc Natl Acad Sci USA* 1990; 87:6009-6613.
44. Martin, D.L., et al., *N Engl J Med* 1990; 323:1161-1167.
45. Martin, M.L., et al., *Lancet* 1986; ii:1043.
46. McNamara, A.M., personal communication.
47. Montenegro, M.A., et al., *J Clin Microbiol* 1990; 28:1417-1421.
48. Moon, H.W., et al., *Am J Clin Nutrition* 1979; 32:119-127.
49. Morgan, R.L., et al., *Infect Immun* 1978; 22:771-777.
50. Morris, J.A.,et al., *J Med Microbiol* 1980; 13:265-271.
51. Oudega, B., et al., *Antonie van Leeuwenhoek* 1988; 54:285-299.

52. Pai C.H., et al., *J Infect Dis* 1988;157:1054-1057.
53. Pai, C.H., et al., *Infect Immun* 1986; 51:16-23.
54. Pararuchuri, D.K., et al., *Proc Natl Acad Sci USA* 1990; 87:333-337.
55. Pecha, B., et al., *J Clin Invest* 1989; 83:2102-2108.
56. Raleigh, E.A., et al., *Nucl Acid Res* 1988; 16:1563-75.
57. Ratnam, S., et al., *J Clin Microbiol* 1988; 26:2006-2012.
58. Riley, L.W., et al., *Infect Immun* 1987; 55:2052-2056.
59. Ritchie M., et al., *J Clin Microbiol* 1992:30;461-464.
60. Rowe, P.C., et al., *Epidemiol Infect* 1993; 110:9-16.
61. Runnels, P.L., et al., *Infect Immun* 55:555-558, 1987.
62. Samadpour M., et al., *Appl Environ Microbiol* 1994; in press.
63. Sancar, A., et al., *J Bacteriol* 1979; 137:692-693.
64. Sherman, P., et al., *Infect Immun* 1991; 59:890-899.
65. Sherman, P.M., et al., *J Med Microbiol* 1988; 26:11-17.
66. Sherman, P., et al., *Infect Immun* 1988; 56:756-761.
67. Smith, D.B., et al., *Gene* 1988; 67:31-40.
68. Sojka, W.J., et al., *J Med Microbiol* 1978. 11:493-499.
69. Swerdlow, D.L., et al., *Ann Intern Med* 1992; 117:812-819.
70. Tarr, P.I., et al., *Am J Epidemiol* 1989; 129:582-586.
71. Tarr, P.I., et al., *J Infect Dis* 1989; 159:344-347.
72. Tarr, P.I., et al., *J Infect Dis* 1990; 162:553-556.
73. Taylor, R.K., et al., *J Bacteriol* 1989; 171:1870-1878.
74. Taylor, R.K., et al., *Proc Natl Acad Sci USA* 1987; 84:2833-2837.
75. Toth, I., et al., *Infect Immun* 1990; 58:1223-1231.
76. Wadolkowski, E.A., et al., *Infect Immun* 1990; 58:2438-2445.
77. Wahl, G.M., et al., *Proc Natl Acad Sci USA* 1987; 84:2160-2164.
78. Wells, J.G., et al., *J Clin Micro* 1983; 18:512-520.
79. Wells, J.G., et al., *J Clin Microbiol* 1991; 29:985-989.
80. Wessels, M.R., et al., *Proc Natl Acad Sci* USA 1991; 88:8317-8321.
81. Yokoyama, H., et al., *Infect Immun* 1992; 60:998-1007.
82. Yu, J., et al., *Mol Microbiol* 1992; 6:411-7.

While the preferred embodiment of the invention has been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.

SEQ ID NO: 1

DNASIS
a pEAR complete


```
        10         20         30         40         50
GGTACCTGTC GCCAGTTCTC CGATCTGTTT ACCCGGGAAA TTATCTCCTA

        60         70         80         90        100
CAGCTTGTCA GAAAGGTCGG TGATGGAGCA TCGNTAATAC GATGCTATAC

       110        120        130        140        150
GATGNATTCA CAGTGCCCGG CCAGAGGATG CCCCGTCGCT GCATATGGAT

       160        170        180        190        200
CAGNGTTGGC AATATCGAAT TGCAGGCTAT AGGCAAAGTT AANGCCCCAT

       210        220        230        240        250
GGAGTAGCAC AAAATATGCC GCACANAGGA AACGGTCTGA ATAACGCAGT

       260        270        280        290        300
GATGAAGAAC TTCTTCAGCA CACTGCTAAA ACGCATAGTG ATCGAGCGCT

       310        320        330        340        350
GATTCTGGCG AACAACTGAA CTAATACATC AGAATCTGCA TTATGTTAAA

       360        370        380        390        400
TAAATATAAA AAGATGGTTT AAATACCCCG TTACTTGTGA CTTACACTAT

       410        420        430        440        450
ACGGTATCGC ATCGTTTAAT ATTCGCACCG GCCAGATTTT TATTTCTATT

       460        470        480        490        500
AGTTGTCACA ATACTGAATG CGTACGACCA CAGTATTCTG GCTCCTGTGT

       510        520        530        540        550
GGTTATGCTT TAATTCTGCG TTCCGGGCAG ATAAGCAGTT GCTTGCAGGA

       560        570        580        590        600
ATCCTTCTTG TGTTAATGTC AGTTCCCCTT TTACCAGTGC TGATTTCCAC

       610        620        630        640        650
ATTCCGTCCA ACAGAGCTTA TAGCCTTTCC CTGGATTATA GCATTGTCCG

       660        670        680        690        700
GCTGAAGTTC TTTTTGAATA ATAATAGAAG CACTGCTGGC AGATCCAGTC

       710        720        730        740        750
CGTTTTTCAT AACCCACTGT ACTGATAACC ATAATCTAAT CAGTAGAAAT

       760        770        780        790        800
TGAGTCGAAA ATAAGCACTA CTCCATACAG GATAATTAGA GGTCAGTTTG

       810        820        830        840        850
ATTATTCACA ATTCATCATC AGCATTTTCT ATTTCTGACG AAATCAATAT

       860        870        880        890        900
GAAAATAACC ATATATGATA ATTATTATAA TAACGGCTTT AATTGGAATA

       910        920        930        940        950
CATATATTAC AACGTATTAT ATATAATTGG TATTCTGGGA ACTATATTCT

       960        970        980        990       1000
CAAAATACAG TAGAAACGAG GTATGTTTCT GGTGGAAAGG ACAGTGGGAT

      1010       1020       1030       1040       1050
TAAAAAGTAA GAATTGATAA AAAAACGCCA GCAACACATA GTGCCGGNGG
```

```
      1060       1070       1080       1090       1100
AGGGAATACC CCATGGAGAA AATGTGATGC CTAGAGCATC ATGANTATAT

      1110       1120       1130       1140       1150
AATTAAAAAT AGTTAGCGTT GTCACTACTT CNACAAAAAT AATTTTCGTA

      1160       1170       1180       1190       1200
GTATAGAAAG ATATTTTTAT GCATGACCTA CCTGAATTTG CTCCGGGTAG

      1210       1220       1230       1240       1250
AGGTTATAAA TAAAAATTGA ATCACGACAA ACACAATATT CATAGTATGG

      1260       1270       1280       1290       1300
CGATGCCTAC GCCAGCAAGA ATAGCGNCAA TAATATTGGG AATATAATAG

      1310       1320       1330       1340       1350
ACACCAGACG CACAGGCATC TCACTCCTTA ACAAACAACA ATCAGGATAT

      1360       1370       1380       1390       1400
TTACTTTTAC CAAGCTAACT GTTTACACCC AAAGTACACA CAATTAACCA

      1410       1420       1430       1440       1450
TTCAATAACA AATGNCAATA TCCATAGCCA TACGACTTTA CCTTGTAATG

      1460       1470       1480       1490       1500
TTCGGTATTT CTTTATAATT ATTCTGGGAA ATCTAACATT TATTTTTAAA

      1510       1520       1530       1540       1550
ATCAAATTAT CTTGTTGTTT AAAAATAAGT TCACATACTT TATCATCTTC

      1560       1570       1580       1590       1600
TTCCGCCATC AACATCTCTG CATACTTAAA CATTTCAGAA CGTTCCTTTA

      1610       1620       1630       1640       1650
GCACAGAAGA GTAATTATAT GTCCAGTTCC CAGGCAATAA TGCTTATGGA

      1660       1670       1680       1690       1700
TATTTAATTC ATAATTTAGA GAATATTTTG CAATAATATT TGGCAGTATT

      1710       1720       1730       1740       1750
CAGAAATACC TGAAAAATCA TACTATACAG CCCTAGGGAA TGGATAAAGA

      1760       1770       1780       1790       1800
TTCTAACAAA GCATTTCAAC AATATATACT TGTTAAAAAT CCATATCGAA

      1810       1820       1830       1840       1850
ATGCCGTTGC AGCATTAATA TATGCCTCAT TCATAAAATG TAAAAGAGCA

      1860       1870       1880       1890       1900
AGCTCGTACC AGTAGGGGGG AAATTCATTG ACATGTCCTG TCAATACGTA

      1910       1920       1930       1940       1950
CAGAGCCCTG CCATGCTTGC CAGCACGTAA CAATTCCGCC CCCAGTAACC

      1960       1970       1980       1990       2000
AGCGTGTTCC CTGATTATTC TCAGGGTTAT ATGTTAGTAT TTTATCAATG

      2010       2020       2030       2040       2050
AGTATAACGG CATCCTGATG ACGCTGTAAA TGAACATTCG CCAGAATTGC

      2060       2070       2080       2090       2100
AGCATAAAGA GCGCAAATGC ATATGCCAGA TGAGCGTGAA TATCAATGAT
```

DNASIS
a pEAR complete


```
      2110        2120        2130        2140        2150
GTCAGGAGCA TGTCAGGCAA GACGCCTGAG TTTGTTGACG TAATTTGTTT

      2160        2170        2180        2190        2200
CGTTTATTTC ATCACTATCG TATGCGTCAA GAGTTTTTTC AAACTCATCC

      2210        2220        2230        2240        2250
CGAAATGGGC CAAGCCTGGT TTCAGGAAAT AAGAAATACC CTTGGTTATT

      2260        2270        2280        2290        2300
GTTCACTTAA AAGTATAATT TGTAATTCAA ATCCTGAGCC TGTAAAGCGG

      2310        2320        2330        2340        2350
GGAGTAACAT ATTCTATTCT GAAGAGAATA AAAGTCGTGA TGCGATGTAT

      2360        2370        2380        2390        2400
CAAGCCCGGA TTGTAATCCC AGATTAACAT AGATCACAAA ACAACTTATT

      2410        2420        2430        2440        2450
CTTCACTAAC GTCAAGATAA ATATCGTTGT ATGCCTTATC ACGACTACGA

      2460        2470        2480        2490        2500
ATACCAGCAA GATACAACTG ACATACGGAA AGATACCACG TTTTTTACTC

      2510        2520        2530        2540        2550
CCGAAAATAA CGCTAAAAAG CTACTTCCCC ATCGTTTGTC CTTAGTATTG

      2560        2570        2580        2590        2600
CCAGCGCCAA CAATGTGGGC TGACATGATA AAGCTGTCTA GGAAATTGTT

      2610        2620        2630        2640        2650
CGCCTCCTCA GCGGACAATC CAAATGGTGA TTGTCTCTGT TAAACGTTTA

      2660        2670        2680        2690        2700
TTTTGAAGGT CGACTGAATA AGGTGATGAC GCTGTAGAAT TTTTCACGTG

      2710        2720        2730        2740        2750
CCACAGAATT TTGAACGCTT TCTCTTACAA TATTTCAATG TTTCTATCAG

      2760        2770        2780        2790        2800
TATTCGCCGG AAGAAGTCAT CGACCAAATC ATCCCAGTCG TCTCGCATCA

      2810        2820        2830        2840        2850
CTGACCATTC ATGGTTGACA TGTGGTGGAA ATCCGCTTCT ACAGTAACCA

      2860        2870        2880        2890        2900
TTTTTTATTC GCAAAACCGA ACACGCCATA CGGATAGCTG TTAACTGGCA

      2910        2920        2930        2940        2950
TGCCGAGGCA GTCGTTATTT ATATTTGGTT TTGTCAATAA TCTTTATTTT

      2960        2970        2980        2990        3000
TTGTAAAAGG CAAATATAAA TTATTCTCAT TATTGTTTGT ATTTGTGTAT

      3010        3020        3030        3040        3050
TGTCTTGCCG GTTAACATGA TCGGAGATTA GTAATATGCG AATAACCACT

      3060        3070        3080        3090        3100
CTGGCTTCCG TAGTCATTCC CTGTCTCGGA TTTTCAGCCA GCAGCATAGC

      3110        3120        3130        3140        3150
TGCTGCAGAG GATGTGATGA TTGTCTCGGC ATCCGGCTAT GAGAAAAAGC
```

(H) = SEQ ID NO:11

(F) = SEQ ID NO:9

(C) = SEQ ID NO:4

DNASIS
a pEAR complete


```
      3160       3170       3180       3190       3200
TGACTAACGC AGCCGCCAGT GTTTCTGTGA TTAGCCAGGA GGAATTGCAG

      3210       3220       3230       3240       3250
TCCAGCCAGT ACCACGATCT GGCGGAGGCT CTGAGATCAG TAGAGGGTGT

      3260       3270       3280       3290       3300
GGATGTTGAA AGTGGTACGG GTAAAACCGG AGGGCTGGAA ATCAGCATCC

      3310       3320       3330       3340       3350
GAGGAATGCC AGCCAGTTAC ACGCTGATAC TGATTGATGG TGTTCGTCAG

      3360       3370       3380       3390       3400
GGCGGAAGCA GTGACGTGAC TCCCAACGGT TTTTCTGCCA TGAATACCGG

      3410       3420       3430       3440       3450
GTTCATGCCC CCTCTGGCCG CCATTGAGCG TATTGAGGTT ATCAGGGGGC

      3460       3470       3480       3490       3500
CGATGTCCAC ACTGTATGGC TCTGATGCGA TGGGCGGTGT GGTGAATATC

      3510       3520       3530       3540       3550
ATTACCAGAA AGAATGCAGA CAAATGGCTC TCTTCCGTCA ATGCAGGGCT

      3560       3570       3580       3590       3600
GAATCTGCAG GAAAGCAACA AATGGGGTAA CAGCAGCCAG TTTAATTTCT

      3610       3620       3630       3640       3650
GGAGCAGTGG TCCCCTTGTG GATGATTCTG TCAGCCTGCA GGTACGCGGT

      3660       3670       3680       3690       3700
AGCACACAAC AGCGTCAGGG TTCATCGGTC ACATCACTGA GCGATACAGC

      3710       3720       3730       3740       3750
AGGCACGCGT ATTCCTTATC CCACGGAGTC ACAGAATTAT AATCTTGGTG

      3760       3770       3780       3790       3800
CACGTCTTGA CTGGAAGGCG TCGGAGCAGG ATGTGCTCTG GTTTGATATG

      3810       3820       3830       3840       3850
GATACCACCC GGCAGCGTTA TGATAACCGG GATGGGCAAC TGGGGAGTCT

      3860       3870       3880       3890       3900
GACGGGGGGA TATGACCGGA CCCTGCGCTA TGAGCGAAAC AAAATTTCAG

      3910       3920       3930       3940       3950
CTGGCTATGA TCATACTTTC ACCTTCGGAA CATGGAAATC GTATCTGAAC

      3960       3970       3980       3990       4000
TGGAACGAGA CAGAAAATAA AGGTCGTGAG CTTGTACGCA GTGTACTGAA

      4010       4020       4030       4040       4050
GCGCGACAAA TGGGGGCTTG CCGGTCAGCC GCGGGAGCTT AAGGAATCGA

      4060       4070       4080       4090       4100
ACCTTATCCT GAATTCATTA CTGCTTACCC CTCTGGGAGA ATCTCATCTG

      4110       4120       4130       4140       4150
GTTACGGTGG GGGGCGAGTT TCAGAGCTCG TCCATGAAAG ACGGAGTTGT

      4160       4170       4180       4190       4200
CCTTGCCAGC ACAGGTGAAA CTTTCCGGCA GAAAAGCTGG TCGGTATTTG
```

D

E

(D) + (E) = regions of TnphoA insertions described in Example 4

DNASIS
a pEAR complete


```
     4210       4220       4230       4240       4250
CTGAGGATGA GTGGCATCTC ACGGATGCAC TTGCGCTGAC TGCGGGCAGC

     4260       4270       4280       4290       4300
CGCTATGAAC ATCATGAGCA ATTCGGGGGA CACTTCAGTC CGCGTGCATA

     4310       4320       4330       4340       4350
TCTGGTCTGG GATGTGGCAG ATGCCTGGAC GCTGAAAGGC GGTGTGACCA

     4360       4370       4380       4390       4400
CGGGATATAA GGCACCCAGA ATGGGGCAGC TACATAAAGG GATTAGTGGT

     4410       4420       4430       4440       4450
GTGTCCGGGC AGGGAAAAAC AAATCTACTT GGTAACCCCG ACCTGAAGCC

     4460       4470       4480       4490       4500
GGAAGAGAGC GTCAGTTATG AGGCTGGGGT GTATTACGAT AACCCCGCCG

     4510       4520       4530       4540       4550
GTCTGAATGC CAATGTCACA GGTTTTATGA CTGACTTCTC CAACAAGATT

     4560       4570       4580       4590       4600
GTCTCTTATT CCATAAATGA TAACACCAAT AGCTATGTAA ACAGCGGAAA

     4610       4620       4630       4640       4650
GGCCCGGTTG CACGGTGTGG AATTTGCCGG CACATTGCCG CTGTGGTCAG

     4660       4670       4680       4690       4700
AGGATGTCAC GCTGTCACTG AATTACACCT GGACCCGAAG TGAACAACGT

     4710       4720       4730       4740       4750
GATGGTGATA ACAAAGGTGC GCCGCTGAGT TATACCCCTG AACACATGGT

     4760       4770       4780       4790       4800
GAATGCGAAA CTGAACTGGC AGATCACCGA AGAGGTGGCA TCATGGCTGG

     4810       4820       4830       4840       4850
GTGCCCGTTA TCGCGGGAAA ACACCACGTT TCACCCAGAA TTATTCGTCA

     4860       4870       4880       4890       4900
CTGAGCGCTG TACAGAAGAA AGTGTATGAT GAGAAAGGAG AATACCTGAA

     4910       4920       4930       4940       4950
AGCCTGGACG GTGGTGGATG CAGGTCTGTC GTGGAAGATG ACGGATGCCC

     4960       4970       4980       4990       5000
TGACGCTGAA TGCTGCGGTG AATAACCTGC TCAACAAGGA TTACAGTGAC

     5010       5020       5030       5040       5050
GTGAGCCTGT ACAGTGCCGG TAAGAGTACG CTGTATGCCG GTGATTACTT

     5060       5070       5080       5090       5100
CCAGACGGGA TCATCAACAA CAGGATATGT GATACCTGAG CGAAATTACT

     5110       5120       5130       5140       5150
GGATGTCGCT GAACTATCAG TTCTGATAAT AACAAAAGGC TATCACTGAC

     5160       5170       5180       5190       5200
GGTAGAATAC GTTGCCACTG CAACTCCTGG CGGAACAGTG GCAACGTNTT

     5210       5220       5230       5240       5250
AGGTTAAGTG CATTTCCGAT CCGCTAATGA GATTTCGTTA CCAACAACTA
```

(G)

(I)

end of (-)

G = SEQ ID NO: 10

I = SEQ ID NO: 12

```
      5260        5270        5280        5290        5300
ATATCGTCAC AGGAAATGCA CGGATTATTT TTAACTTATC ATTTACATAC

      5310        5320        5330        5340        5350
TTGTCCAGAG TGTNAGCGCA CCGCGACGGA CGTGGGGTAA AAATTAGTTT

      5360        5370        5380        5390        5400
ACAGAGAGAG TGACGTTCCA GGGGAACAAC TCTTTCATGC GGTTGGCAGG

      5410        5420        5430        5440        5450
CCAGGTGTTG GTTACACTGA TCACGTGGGC GTTGGCCACG TTTCCGGNTC

      5460        5470        5480        5490        5500
GATTCCGTTA AGTTTTGGAG CTACCGATCA GGCTGTACAT CACTGNCGCA

      5510        5520        5530        5540        5550
CTATCGCTCG TCATCTCAAA GTCCTGTCTC GTCAGCAGGA AGGTATCATT

      5560        5570        5580        5590        5600
CTCTCCCGCC ATTTTTCCAG GGGNCCGGTC AGATAAGTCC CTTTGTCTAT

      5610        5620        5630        5640        5650
CGCTGACTCC TGACTCATAA CCCGGTTAGC AGAATGCAGG NTCACCACTC

      5660        5670        5680        5690        5700
GCCACGACCA AATCCAAATA AGTCAATTGC ACCTTCTCAA TCGCCATTTT

      5710        5720        5730        5740        5750
GTCAGTAAGC GTACAGCCTC AACTGATGGT ATCTTCACCA TCAATGACAA

      5760        5770        5780        5790        5800
CGGTGATCGC AATTTTACTG ACGTTCGCCG GAACACGATC CAGTGCTATC

      5810        5820        5830        5840        5850
TCAATGCTGG CCTGCTGCGA ACCGGTAACG AGCCTGACTG CCCCCTCAGG

      5860        5870        5880        5890        5900
AGAAGACAAA TTATTATAAA AGATAAAGTC AGAATCGCCA CTGACCTTTC

      5910        5920        5930        5940        5950
CCTGAGCATT AAGCATGAAC AGGGAGGTAT CGGGTTCGCC TTAAAAGCCG

      5960        5970        5980        5990        6000
GATTTGNCAG GGTACTGAAG ATTCAGCCTG ATCGNAGATT GCTGAAGGGG

      6010        6020        6030        6040        6050
TATGTATTGT CCGGATTGTA AATTCATATT AACTCTCCTG ATTTNTGATT

      6060        6070        6080        6090        6100
ATTATTAATG CGCAGCGTTT ATATATGTTC CCTAGGCTTA GTTCTGGACG

      6110        6120        6130        6140        6150
CTGGATATTC GGTGAGGCGT AAATATGGTA TGACACCATT TTTCATAACG

      6160        6170        6180        6190        6200
CTGAAGTTTC TATACCTGTT GAATTTGAAT TTTCATTGAC CGGGTATCTT

      6210        6220        6230        6240        6250
ATTTTCCAGG GCCCGTTCCT TCATAAGTCG CAAAAGTAAC ATATATCCGA

      6260        6270        6280        6290        6300
AGGGCATGCT GTTGATATCA GACACGGAAT ACTGGCTTTA ACCAGCAACC
```

DNASIS
a pEAR complete


```
      6310       6320       6330       6340       6350
ACAGATAAAC CCGGGGCCTG CATAAAAGGT TCATGCCAGA ATTAACATAG

      6360       6370       6380       6390       6400
CTCTGCTTTC TGCCCCCCCC TTTCCCATAT TCACCCGTTG ATAGCGGATC

      6410       6420       6430       6440       6450
GATACCCAAA AAAAACCCCG GCATTGCCGG GGTCAGACTC AGCGTCTTTT

      6460       6470       6480       6490       6500
AGACATTGAC GCCGTGCTGG GAAGCAAGCG CTGACAGACC ACCGGCAAAA

      6510       6520       6530       6540       6550
CNCTGGCCAA CAGCTTTAAA CTTCCACTCA GTACCATGGC GATACAGTTC

      6560       6570       6580       6590       6600
ACCGAAGACC ATTGCGGTTT CGGTTGAGGC GTCTTCAGAC AGATCGAAAC

      6610       6620       6630       6640       6650
GGGCAATTTC CGTCCCGTTG TCGTTGTTGT AAACGCGCAT GAAGCTGTTG

      6660       6670       6680       6690       6700
CTCACCATGC CGAAGTTTTG TTTACGCGCT TCTGCATCAT AGATGGTAAC

      6710       6720       6730       6740       6750
GGCAAATACC AGTTTTTTGA TGTCTGCTGA GACTTTGGTC AGATCGATTT

      6760       6770       6780       6790       6800
TGACCTGCTC ATCGTCGCCG TCGCCTTCAC CGGTACGGTT GTCGCCCTGG

      6810       6820       6830       6840       6850
TGCTCTACTG CGCCATCAGG GCTGGTTTTA TTATTGAAGA AAATGAAATG

      6860       6870       6880       6890       6900
GGCATCTGAC AGTACTTTAC CGTCTTCACC TACTGCGAAT ACGGAAGCGT

      6910       6920       6930       6940       6950
CCAGANCAAA ACCCTGACCA TCGGTTACAC GGGCATCCCA GCCCAGGCCA

      6960       6970       6980       6990       7000
ACCATAGCGA CATTCATGGT TGGTGCTTCT TTGGTCAGAG ATACGTTGCC

      7010       7020       7030       7040       7050
GCCTTTTACG AGAGAAACTG CCATTTTTAG CTCCTGCAAA CAGNTGAATG

      7060       7070       7080       7090       7100
AGGCTGAATA ACACCCCCAG AAATGAAAAG TTACTTTTCG ATCAGGACGC

      7110       7120       7130       7140       7150
GTTAATNCCG TACTGAGCAC ATACAGATGC CAGACCACCA GCATAACNCT

      7160       7170       7180       7190       7200
GTNCTACTGC GCGGAATTTC CACTCACCAT TGTGGCGAGA CAGCTCGCCG

      7210       7220       7230       7240       7250
CGCAGCATGG CAGTCTCAGT GGACGCATCT TCGGTCAGAT CGTAGCGAGC

      7260       7270       7280       7290       7300
GACTTCAGTC TGGTTATCGT CATTAACCAG ACGAATAAAC GCACCGGATA

      7310       7320       7330       7340       7350
CCTGACCACA GCTCTGGCGA CGAGCCTGAG CATCGTGGAT GGTCACAACG
```

(A) = SEQ ID NO: 2

end of (A)

end of (B)

DNASIS
a pEAR complete


```
      7360        7370        7380        7390        7400
AAGATGATCT TGTCAACTTC AGACGGGACG GCGTCCAGTT TAATTTTCAG

      7410        7420        7430        7440        7450
CGATTCATCA TCACCATCGC CCTCACCGGT GCGGTTATCG CCGGTGTGCG

      7460        7470        7480        7490        7500
TTACGGAACC GTCGGATGAC GTCAGGTTGT TATAGAAGAT GAAATCTGAA

      7510        7520        7530        7540        7550
TCGCCGCGCA CTTTGCCGTT TGAGGCCAGC AGGAATGCTG AAGCATCCAG

      7560        7570        7580        7590        7600
GTCAAAGTCC TGACCGTCTG TTGAACGCGC ATCCCAGCCA AGGCCCACCA

      7610        7620        7630        7640        7650
GGACATTTTT CATTGACGGA GCTGCTTTAC TCAGGGAGAC GTTCCCGCCT

      7660        7670        7680        7690        7700
GTGGAAAGAG AAACACTCAT AAAATACCCT CTTCGATTAG TAATTGTTCA

      7710        7720        7730        7740        7750
GGTTAACACT TAAGGGGATT ATCTCCCCTT TTCCTCAGAT TCAGGTGTGC

      7760        7770        7780        7790        7800
CCGGGAACAT GACGCTTGCG AGAATGCCCA GCGCCAGTAC ACCCAGTACA

      7810        7820        7830        7840        7850
ACATACAGGC TGGTTGTTGC CGCGATGCTG TAACCATGAT GCCAGATGTG

      7860        7870        7880        7890        7900
ATCGATCGCA TTCAGGCCGA GTTTTGCCAC GATGAAGAAC AGCANCACGA

      7910        7920        7930        7940        7950
TAGCGGCCTT CTCCAGATGN ACCAGGTNCT GTTTCAGTGC CTCNAGGACA

      7960        7970        7980        7990        8000
AAATACAGAG TACGCAGACC CAGGATAGCA AACATCATGG CACTATAGAC

      8010        8020        8030        8040        8050
GATGAAGCGG TTCACGACTG ACGGCAATGA TTTCCGGTAC C.........
```

B

B = SEQ ID NO:3

END OF SEQ ID NO:1

SEQ NO: 5

E.coli O157:H7 adhesin sequence>> 1 MRITTLASVVIP////CLGFSASSI 63
Homology (if any)>> : : : ::::
IrgA amino acid sequence [V.cholera]>> MSRFNPSPVSLSVTLGLMFSAS//

SEQ NO. 6

64 AAAED//////VMIVSASGYEKKLTNAAASVSVISQEELQSSQYHDLAEALRSVEGVDVE 225
: :.: :.:.:.:: . . :: ::.:::::.:.:.: : :. .:::.:: :: :
AFAQDATKTDETMVVTAAGYAQVIQNAPASISVISREDLESRYYRDVTDALKSVPGVTV/

226 SGTGKTGGLEISIRGMPASYTLILIDGVRQGGSSDVTPNGFSAMNTGFMPPLAAIERIEV 405
:: .:::::: ..:::::.:: :: . . :..::: :::::::
//TGGGDTTDISIRGMGSNYTLILVDGKRQTSRQTRPNSDGPGIEQGWLPPLQAIERIEV

406 IRGPMSTLYGSDAMGGVVNIITRKNADKWLSSVNAGLNLQESNKWGNSSQFNFWSSGPLV 585
:::::::::::::.:::.:::::::. .: .: .:::. :. ::. .:::
IRGPMSTLYGSDAIGGVINIITRKDQQQWSGNVQLSTVVQENRASGDEQSANFFVTGPLS

586 DDSVSLQVRGSTQQRQGSSV//TSLSDTAGTRIPYPTESQNYNLGARLDWKASEQDVLWF 759
: ..:::: : : :: . . . : . .:. : ::
D/ALSLQVYGQTTQRDEDEIEHGYGDKSLRSLTSKLNYQLNPDHQLQLEAGVSAQDRENN

760 DMDTTRQRYDNRDGQLGSLTGGYDRTLRYERNKISAGYDGTFTFGTWKSYLNWNETENKG 939
. . : : .: : : .:: . : ::
VGKSAQSSGC//////RGTCSNTDNQYRRNHVAVSHQGDWQGV/GQSDTYLQYEENTNKS

940 RELVRSVLKRDKWGLAGQPRELKESNLILNSLLLTPLGESHLVTVGGEFQSSSMKDGVV 1116
:: : . : :. :.:: : :
RE//////////////////MSIDNTVFKSTLVAPIGEHMLSFGVEGKHESLEDKTSN

1117 LASTGETFRQKSWSVFAEDEWHLTDALALTAGSRYEHHEQFGGHFSPRAYLVWDVADA 1290
:. :. : :::: : . :: : : .: . .: ::::: : ::..
KISSRTHISNTQWAGFIEDEWALAEQFRLTFGGRLDHDKNYGSHFSPRVYGVWNLDPL

1291 WTLKGGVTTGYKAPRMGQLHKGISGVSGQGKTNLLGNPDLKPEESVSYEAGVYYDNPA 1464
::.::::.::..::.. .. ::: : :. :::::.:: :.. : . : .
WTVKGGVSTGFRAPQLREVTPDWGQVSGGG//NIYGNPDLQPETSINKELSLMYSTGS

1465 GLNANVTGFMTDFSNKI//VSYSINDNTNSYVNSGKA///RLHGVEFAG/////TLPL 1608
:: :..: : :: .:: :. : : : :. : :: :
GLAASLTAFHNDFKDKITRVACPANICTAGPNQWGATPTYRVNIDEAETYGAEATLSL

1609 W/SEDVTLSLNYTWTRSEQRDGDNKGAPLSYTPEHMVNAKLNWQITEEVASWLGARYR 1779
.: : :: .:::.: :::. :. : :: :.:. :: :.:: :. . :: ::
PITESVELSSSYTYTHSEQKSGNFAGRPLLQLPKHLFNANLSWQTTDRLNSWANLNYR

Page 10

```
1780 GKTPRFTQNYSSLSAVQKKVYDEKGEYLKAWTVVDAGLSWKMTDALTLNAAVNNLLNK 1953
     ::     :     :                ..: .: :... .::  :. ::: :: ..
     GK///EMQPEGGASNDDFIAP////////SYTFIDTGVTYALTDTATIKAAVYNLFDQ

1954 DYSDVSLYSAGKSTLYAGDYFQTGSSTTGYVIPERNYWMSLNYQF* 2091 END OF SEQ NO:5
     .          ::           :::   : ::. :. . ::
     E////////////VNYAEY/////////GYVEDGRRYWLGLDIAF* END OF SEQ NO:6
```

Slashes (/) indicate areas where gaps were inserted into the amino acid sequence to give the best possible alignment

":" indicates an identical amino acid
"." indicates a conserved amino acid

Page 11

SEQ ID NO: 7

E.coli O157:H7 adhesin sequence>> 1 MRITTLASVVIPCLGFSASSI///// 63
Homology (if any)>> : : :: . .
CirA amino acid sequence [E.coli]>> MFRLNPFVRVGLCLSAISCAWPVLAV

SEQ ID NO: 8

64 AAAEDVMIVSASGYEKKLTNAAASVSVISQEELQSSQYHDLAEALRSVEGVDVESGTGKT 243
. :.:.:: :. : .: ::.:::.::.:: .: . :. : :: . .
DDDGETMVVTASSVEQNLKDAPASISVITQEDLQRKPVQNLKDVLKEVPGVQLTNEGDNR

244 GGLEISIRGMPASYTLILIDGVRQGGSSDVTPNGFSAMNTGFMPPLAAIERIEVIRGPMS 423
:. :::: .::::::.:: : :. .:::::::.:::::
KGV//SIRGLDSSYTLILVDGKRVNSRNAVFRHNDFDLNWI///PVDSIERIEVVRGPMS

424 TLYGSDAMGGVVNIITRKNADKWLSSVNAGLNLQESNKWGNSSQFNFWSSGPLVDDSVSL 603
.:::::::.:::::::.: :: .: .:: :.. :..::::.: . .
SLYGSDALGGVVNIITKKIGQKWSGTVTVDTTIQEHRDRGDTYNGQFFTSGPLIDGVLGM

604 QVRGSTQQRQGSSVTSLSDTAGTRIPYPTESQNYNLGARLDWKASEQDVLWFDMDTTRQR 783
. :: .:. . . : : . . : .. ::
KAYGSLAKREKDDPQNSTTTDTGETPRIEGFSSRDGNVEFAWTPNQNHDFTAGYGFDRQD

784 YDNRDGQLGSLTGGYDRTLRYERNKISAGYDGTFTFGTWKSYLNWNETENKGRELVRSVL 963
:. : :: : .. . .:: . . ::: .
RDSDSLDKN/////////RLERQNYSVSHNGRWDYGTSELKYYGEKVENKN////////

964 KRDKWGLAGQPRELKESNLILNSLLLTPLGESHLVTVGGEFQSSSMKDGVVLASTGETF 1140
::: . : . .:::::.. . : : :
//////PGNSSPITSESNTVDGKYTLPLTAINQFLTVGGEWRHDKLSDAVNLTGGTSSK

1141 RQKS/WSVFAEDEWHLTDALALTAGSRYEHHEQFGGHFSPRAYLVWDVADAWTLKGGV 1311
: ...: :::: . . :::: : : . :: .: :.::::::::.. : :.:::
TSASQYALFVEDEWRIFEPLALTTGVRMDDHETYGEHWSPRAYLVYNATDTVTVKGGW

1312 TTGYKAPRMGQLHKGISGVSGQGKTNLLGNPDLKPEESVSYEAGVYYDNPAGLNANVT 1485
: .::: . :: . : .: ..:.:::::: : :.: :.:: : :
ATAFKAPSLLQLSPDWTSNSCRGACKIVGSPDLKPETSESWELGLYYMGEEGWLEGVE

1486 GFMTDFSNKI/VSYSINDNTNSYVNSGKARLHGVEFAGTLPLWSEDVTLSLNYTWTRS 1656
.: : : . ::. .. : : : ..: .
SSVTVFRNDVKDRISISRTSDVNAAPGYQNFVGFETGANGRRIPVFSYYNVNKARNQG

1657 EQRDGDNKGAPLSYTPEHMVNAKLNWQITEEVASWLGARYRGKTPRFTQNYSSLSAVQ 1830
. . : . .. : :
VETELKIPFNDEWKLSINYTYNDGRDVSNGENKPLSDLPFHLALEDWSFYVSGHYTGQ

```
1831 KKVYDEKGEYLKAWTVVDAGLSWKMTDALTLNAAVNNLLNKDYSDVSLYSAGKSTLYA 2004
     :.      .    .:. . : .:..:   . : : : :: .:          .
     KRADSATAKTPGGYTIWNTGAAWQVTKDVKLRAGVLNLGDK//////////TANGTL

2005 GDYFQTGSSTTGYVIPERNYWMSLNYQF* 2091  END OF SEQ ID NO:7
                :    : :.:...:.:
     DWKPDLSRDDYSYNEDGRRYFMAVDYRF  END OF SEQ ID NO:8
```

Slashes (//) indicate areas where gaps were inserted into the amino acid sequence to give the best possible alignment

":" indicates an identical amino acid
"." indicates a conserved amino acid

DNASIS
Untitled1


```
        10         20         30         40         50         60
ATGGAAGCAG CAAATTTAAG TCCTTCTGGT GCAGTAATGC CGCTGGCGAC CTCACTCAGT

        70         80         90        100        110        120
GGAAATAACT CAGTGGATGA GAAGACAGGA GTGATTAAAC CAGAAAATGG AACAAATCGC

       130        140        150        160        170        180
ACCGTTAGAG TTATAGCCGG ATTAGCACTT ACCACTACGG CTCTGGCAGC TCTAGGTACA

       190        200        210        220        230        240
GGTATTGCAG CGGCATGCTC GGAGACGAGC AGCACAGAAT ACTTAGCCCT GGGTATTACT

       250        260        270        280        290        300
TCTGGCGTAC TAGGTACTCT TACTGCGGTT GGCGGTGCAT TAGCGATGAA ATATGCCTAA

       310        320        330        340        350        360
.......... .......... .......... .......... .......... ..........
```